Variable and Fixed Interval Exponential Smoothing

Javier R. Movellan
University of California San Diego

1 Introduction

Exponential smoothers provide an efficient way to compute moving averages of signals Tong [1990]; Box et al. [2004]; Lutkepohl [2010]. This can be particularly useful for real time applications. Here we define, describe and analyze exponential smoothing algorithms, with an eye towards practical applications.

Let $X_{t_1}, X_{t_2}, \cdots$ be a sequence of random variables observed at times $t_1 \leq t_2 \leq t_3 \leq \cdots$ be the times at which the variables are observed. We define the exponentially smoothed average $\hat{X}_t$ evaluated at time t as follows

$$\hat{X}_t = \frac{\tilde{X}_t}{\tilde{w}_t} \tag{1}$$

$$\tilde{X}_t = \sum_{t_k \leq t} e^{-(t-t_k)/\tau} X_k \tag{2}$$

$$\tilde{w}_t = \sum_{t_k \leq t} e^{-(t-t_k)/\tau} \tag{3}$$

where $\tau \geq 0$ is known as the time scale parameter.

2 Recursive Update Equations

A key advantage of exponential smoothing is that it does not require storage of past observations in memory. The current smoothed average, the current new observation, the time of the last observation and the time of the current observation determine the new smoothed value.

Version 1: Note

$$\tilde{X}_{t_k} = X_{t_k} + e^{-\delta(k,k-1)/\tau} X_{t_{k-1}} + e^{-\delta(k,k-2)/\tau} X_{t_{k-2}} + \cdots + e^{-\delta(k,1)/\tau} X_{t_1} \tag{4}$$

where $\delta(k,j) = t_k - t_j$. Moreover

$$\tilde{X}_{t_{k-1}} = X_{t_{k-1}} + e^{-\delta(k-1,k-2)/\tau} X_{t_{k-2}} + \cdots + e^{-\delta(k-1,1)/\tau} X_{t_1} \tag{5}$$

Thus

$$\tilde{X}_{t_k} = X_{t_k} + \alpha_{t_k} \tilde{X}_{t_{k-1}} \tag{6}$$

$$\alpha_{t_k} = e^{-\delta(k,k-1)/\tau} \tag{7}$$

with initial condition $\hat{X}_{t_1} = X_1$. Similarly

$$\tilde{w}_{t_k} = \tilde{w}_{t_k} + \alpha_{t_k} \tag{8}$$

with initial condition $\tilde{w}_{t_1} = 1$. If t is not one of the sampling times, i.e., $t \in (t_k, t_{k+1})$ then

$$\bar{X}_t = e^{-(t-t_k)/\tau} \tilde{X}_{t_k} \tag{9}$$

$$\tilde{w}_t = e^{-(t-t_k)/\tau} \tilde{w}_{t_k} \tag{10}$$

$$\hat{X}_t = \frac{\tilde{X}_t}{\tilde{w}_t} = \hat{X}_{t_k} \tag{11}$$

Version 2; We shall see later that the previous update equation may run into numerical problems as $\alpha \to 0$. Here we see a version of the update equations that avoids this problem, though it runs into other issues as $\alpha \to 1$. The trick is to multiply the numerator and denominator of (1) by $1 - \alpha$, which is acceptable as long as $\alpha < 1$. Regarding the numerator of (1), let

$$\begin{aligned}
\bar{X}_{t_k} &= (1 - \alpha_{t_k})\tilde{X}_{t_k} = (1 - \alpha_{t_k})X_{t_k} \\
&\quad + (1 - \alpha_{t_k})\alpha_{t_k}\tilde{X}_{t_{k-1}} \\
&= (1 - \alpha_{t_k})X_{t_k} + \alpha_{t_k}\bar{X}_{t_{k-1}} \tag{12} \\
&= X_{t_k} + (1 - \alpha_{t_k})(X_{t_k} - \bar{X}_{t_{k-1}}) \tag{13}
\end{aligned}$$

Regarding the denominator of (1) let

$$\begin{aligned}
\bar{w}_{t_k} &= (1 - \alpha_{t_k})\tilde{w}_{t_k} = (1 - \alpha_{t_k}) + (1 - \alpha_{t_k})\alpha_{t_k}\tilde{w}_{t_{k-1}} \\
&= (1 - \alpha_{t_k}) + \alpha_{t_k}\bar{w}_{t_{k-1}} \\
&= 1 + (1 - \alpha_{t_k})(1 - \bar{w}_{t_{k-1}}) \tag{14}
\end{aligned}$$

Then

$$\hat{X}_{t_k} = \frac{\tilde{X}_{t_k}}{\tilde{w}_{t_k}} = \frac{(1 - \alpha_{t_k})\tilde{X}_t}{(1 - \alpha_{t_k})\tilde{w}_t} = \frac{\bar{X}_{t_k}}{\bar{w}_{t_k}} \tag{15}$$

3 Constant Sampling Rate

It is useful to analyze how the update equations behave when the sampling rate is constant. In this case the time interval between observations is a constant δ and thus $\alpha = e^{-\delta/\tau}$ is also constant.

Update Equations Version 1: Consider update equations of the form

$$u_{k+1} = c + \alpha u_k \tag{16}$$

where c is a constant. At equilibrium

$$u_k = c + \alpha u_k \tag{17}$$

$$u_k = \frac{c}{1 - \alpha} \tag{18}$$

Thus at equilibrium

$$\tilde{w}_{t_k} = 1 + \alpha\tilde{w}_{t_{k-1}} \tag{19}$$

$$\tilde{w}_{t_k} = \frac{1}{1 - \alpha} \tag{20}$$

Note that w_{t_k} can grow very large for $\alpha \to 1$, potentially causing numerical problems. If the X_{t_k} variables are independent identically distributed random variables with mean μ and variance σ^2 then the equilibrium equations for the expected value are as follow:

$$E[\tilde{X}_{t_k}] = \mu + \alpha E[X_{t_k}] \tag{21}$$

$$E[\tilde{X}_{t_k}] = \frac{\mu}{1 - \alpha} \tag{22}$$

$$E[\hat{X}_{t_k}] = \frac{E[\bar{X}_{t_k}]}{\bar{w}_{t_k}} = \mu \tag{23}$$

and the equilibrium equations for the variance are as follow

$$Var[\tilde{X}_{t_k}] = \sigma^2 + \alpha^2 Var[X_{t_k}] \tag{24}$$

$$Var[\tilde{X}_{t_k}] = \frac{\sigma^2}{1 - \alpha^2} \tag{25}$$

$$Var[\hat{X}_{t_k}] = \frac{(1-\alpha)^2}{1-\alpha^2}\sigma^2 = \frac{1-\alpha}{1+\alpha}\sigma^2 \tag{26}$$

Update equations Version 2: In this case

$$\bar{X}_{t_k} = (1-\alpha)X_{t_k} + \alpha\bar{X}_{t_{k-1}} = X_{t_k} + (1-\alpha)(X_{t_k} - \bar{X}_{t_{k-1}}) \tag{27}$$

$$\bar{w}_{t_k} = (1-\alpha) + \alpha\bar{w}_{t_{k-1}} = 1 + (1-\alpha)(1 - \bar{w}_{t_{k-1}}) \tag{28}$$

The equilibrium equation for $\bar{w}$ is as follows

$$\bar{w}_{t_k} = (1-\alpha) + \alpha\bar{w}_{t_k} \tag{29}$$

$$\tag{30}$$

Thus at equilibrium

$$\bar{w}_{t_k} = 1 \tag{31}$$

This can be also seen using the sum of geometric series

$$\alpha^0 + \alpha^1 + \alpha^2 + \cdots + \alpha^{n-1} = \frac{1 - \alpha^n}{1 - \alpha} \tag{32}$$

Thus

$$\bar{w}_{t_k} = 1 - \alpha^k \tag{33}$$

$$\lim_{k \to \infty} \bar{w}_{t_k} = 1 \tag{34}$$

Thus, asymptotically, as $t_k \to \infty$

$$\hat{X}_{t_k} = \alpha X_{t_k} + (1-\alpha)\bar{X}_{t_{k-1}} = X_{t_k} + (1-\alpha)(X_{t_k} - \hat{X}_{t_{k-1}}) \tag{35}$$

If the observations are independent identically distributed with mean μ and variance σ^2 then the asymptotic equations for the mean and variance of the smoother are as follow

$$E[\hat{X}_t] = \alpha\mu + (1-\alpha)E[\hat{X}_t] \tag{36}$$

$$Var[\hat{X}_t] = \alpha^2\sigma^2 + (1-\alpha^2)Var[\hat{X}_t] \tag{37}$$

Thus, asymptotically, as $t_k \to \infty$

$$E[\hat{X}_{t_k}] = \frac{\alpha\mu}{\alpha} = \mu \tag{38}$$

$$Var[\hat{X}_{t_k}] = \frac{(1-\alpha)^2}{1-\alpha^2}\sigma^2 = \frac{1-\alpha}{1+\alpha}\sigma^2 \tag{39}$$

Effective Number of Averaged Observations: If $\hat{X}$ were the average of n independent observations the variance of $\hat{X}$ would be σ^2/n. Thus asymptotically, the effective number of observations averaged by the exponential smoother can be defined as follows

$$n = \frac{1+\alpha}{1-\alpha} \tag{40}$$

$$\alpha = \frac{n-1}{n+1} \tag{41}$$

Thus, for $\alpha = 0$ we average one observation. The number of averaged observations increases unboundedly as α increases.

Effective Time Window: In a time window of size T we get T/δ observations. Thus

$$\alpha = e^{-\delta/\tau} = \frac{T/\delta - 1}{T/\delta + 1} = \frac{T - \delta}{T + \delta} \tag{42}$$

Taking logs

$$-\frac{\delta}{\tau} = \log(T - \delta) - \log(T + \delta) \tag{43}$$

$$\frac{1}{\tau} = \log(T + \delta) - \log(T - \delta) \tag{44}$$

$$\frac{1}{\tau} = \frac{\log(T + \delta) - \log(T - \delta)}{\delta} \tag{45}$$

Thus, in the limit as $\delta \to 0$

$$\frac{1}{\tau} = 2\frac{d\log(T)}{dT} = \frac{2}{T} \tag{46}$$

$$\tau = \frac{T}{2} \tag{47}$$

i.e., the effective window size T is twice the time scale parameter τ. Turns out for $\delta < 0.1$, the approximation

$$\frac{\log(T + \delta) - \log(T - \delta)}{\delta} \approx \frac{2}{T} \tag{48}$$

is quite good. Under this approximation we get

$$\delta/\tau = \frac{2\delta}{T} = \frac{2}{n} \tag{49}$$

$$\alpha = e^{-\delta/\tau} = e^{-2/n} \tag{50}$$

$$n = 2\log(\alpha^{-1}) \tag{51}$$

4 Alternative Approaches

Kalman filters provide an alternative way to deal with variable interval observations. In the constant interval case Kalman filters and exponential smoothers are asymptotically equivalent. In the variable interval case they both weight less the last smooth value the longer the time between the last observation and the current observation. However they differ in the way they do so. An advantage of exponential smoothers over Kalman filters is their simplicity. An advantage of Kalman filters is that they provide estimates of the uncertainty of the smoothed values. This can be useful in some situations.

References

G. E. P. Box, G. M. Jenkins, and Reinsel G. C. *Time Series Analysis: Forecasting and Control.* Wiley, 2004.

H Lutkepohl. *New Introduction to Multiple Time Series Analysis.* Springer, 2010.

H. Tong. *Non-linear Time Series: A Dynamical System Approach.* Oxford University Press., 1990.